DTU Library

Exploring the Use of Design of Experiments in Industrial Processes Operating Under Closed-Loop Control

Experimentation in closed-loop systems

Capaci, Francesca; Kulahci, Murat; Vanhatalo, Erik

Published in:
Quality and Reliability Engineering International

Link to article, DOI:
10.1002/qre.2128

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Research Article

Quality and Reliability Engineering International

(wileyonlinelibrary.com) DOI: 10.1002/qre.2128

Published online 6 February 2017 in Wiley Online Library

Exploring the Use of Design of Experiments in Industrial Processes Operating Under Closed-Loop Control

Francesca Capaci,[a*†] Bjarne Bergquist,[a] Murat Kulahci[a,b] and Erik Vanhatalo[a]

Industrial manufacturing processes often operate under closed-loop control, where automation aims to keep important process variables at their set-points. In process industries such as pulp, paper, chemical and steel plants, it is often hard to find production processes operating in open loop. Instead, closed-loop control systems will actively attempt to minimize the impact of process disturbances. However, we argue that an implicit assumption in most experimental investigations is that the studied system is open loop, allowing the experimental factors to freely affect the important system responses. This scenario is typically not found in process industries. The purpose of this article is therefore to explore issues of experimental design and analysis in processes operating under closed-loop control and to illustrate how Design of Experiments can help in improving and optimizing such processes. The Tennessee Eastman challenge process simulator is used as a test-bed to highlight two experimental scenarios. The first scenario explores the impact of experimental factors that may be considered as disturbances in the closed-loop system. The second scenario exemplifies a screening design using the set-points of controllers as experimental factors. We provide examples of how to analyze the two scenarios. © 2017 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd

Keywords: Design of Experiments; engineering control; feedback adjustment; simulation; Tennessee Eastman process

1. Introduction

Industrial processes often involve automated control systems to reduce variation of quality characteristics or variables affecting plant safety. Sometimes, the control relies on human intervention, such as subjective evaluation of the process state followed by an operator's control action. Processes operating under such control regimes are operating under some form of closed-loop control. Experimenting in these processes will be challenging due to controllers' continuous interference, see Box and MacGregor.[1,2] Because the control action will potentially eliminate the impact of experimental factor changes, experimentation in closed-loop systems may be seen as futile. However, we argue that well designed and properly analyzed experiments run under such conditions can yield valuable information.

This article relates to system identification, which aims at building mathematical models of dynamic systems based on observed data from the system, see Ljung.[3] Experimental design in that sense typically concerns the selection of a proper input signal disturbance to discover the causal relationships between the disturbance and the responses or manipulated variables. This way, system identification allows for the estimation of model parameters to optimize a feedback controller, see, e.g. Jansson.[4] Typically, experimental design research in the system identification field studies 'optimal' input signals to model the system.

In this article, we are primarily concerned with factor screening, factor characterization or process improvement and optimization rather than modeling process dynamics through factors that are already known to affect the response. Similar to system identification experiments, allowable factor ranges are usually restricted, the experiments could be run in full-scale production and the number of experimental runs are limited. However, compared to system identification, the experiments we consider are run for longer periods of time and, most importantly, they have a more overarching purpose of improving or optimizing a process rather than to guarantee stability of a control loop.

[a]*Luleå University of Technology, Luleå, Sweden*
[b]*Technical University of Denmark, Kongens Lyngby, Denmark*
**Correspondence to: Francesca Capaci, Luleå University of Technology, Luleå, Sweden.*
[†]*E-mail: francesca.capaci@ltu.se*

Closed-loop environments add complexity to experimental design and analysis because the control strategy affects the choice of experimental factors. For example, some input variables are manipulated within control loops and therefore may not be suitable as experimental factors. Moreover, even though closed-loop operation is common, we argue that Design of Experiments (DoE) literature typically rests on the implicit assumption that the studied system is operating in open loop, hence allowing the experimental factors to freely affect the response(s). However, as pointed out by, e.g. Vanhatalo and Bergquist[5] and Hild *et al.*,[6] process control systems are designed to maintain the important process variables at their set-points with low variability. Hence, control loops may counteract deliberate changes of experimental factors and thereby displace the effect from typical responses to manipulated variables. An analysis implication is that these manipulated variables instead may have to be used as responses to understand the experimental factors' impact on the system.

The purpose of this article is therefore to explore experimental design and analysis issues in processes operating under closed-loop control and to illustrate how DoE can add value in improving or optimizing such processes. We will pursue this through the help of a process simulator. Process simulators in general have limitations in mimicking the behavior of a real process, but they also offer the flexibility required for methodological developments without jeopardizing plant safety or product quality.

A well-known simulator in the engineering control community is the Tennessee Eastman (TE) challenge chemical process simulator first described by Downs and Vogel.[7] The TE simulator has been primarily used in the development of different process control strategies and for the development of statistical process monitoring methods mainly in chemometrics literature, see for example Kruger *et al.*[8] In this article, we run the TE process with a decentralized control strategy to simulate and illustrate experiments in a closed-loop system.

The remainder of this article is organized as follows: Section 2 establishes important concepts and provides a general comparison of open loop and closed-loop systems from a DoE perspective. Section 3 provides a general description of the TE process simulator and the chosen control strategy. Section 3 also outlines the two experimental scenarios we illustrate in closed-loop operation of the process. The experimental scenarios are elaborated and analyzed in Sections 4 and 5, respectively. Finally, conclusions and discussion are provided in Section 6.

2. Experiments run in open vs. closed-loop systems

Experiments imply that one or many input variables (experimental factors) are allowed to vary to affect the output (response(s)) with the aim of revealing potential causal relationships (effects) between factors and responses, and providing estimates of these effects. The response could be also affected by random disturbances, see Figure 1.

In a process operating under closed-loop control, unwanted variable deviations are mitigated by adjusting a manipulated variable, see Figure 2.

From an experimental perspective, the manipulated variables involved in control loops are not potential experimental factors. In fact, because manipulated variables are involved in control loops, the control engineers have an idea, e.g., from a past experiment, how the manipulated variables affect the response. In relation to Figure 2, the experimental factors in a closed-loop setting should be viewed as disturbances to the system operating under closed-loop control. The potential effects of a disturbance on the controlled variable(s) are therefore typically masked and displaced to one or several manipulated variables if the control system is working

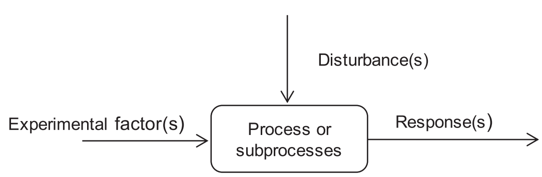

Figure 1. Experimental paradigm for open-loop operation. Figure inspired by Montgomery.[9]

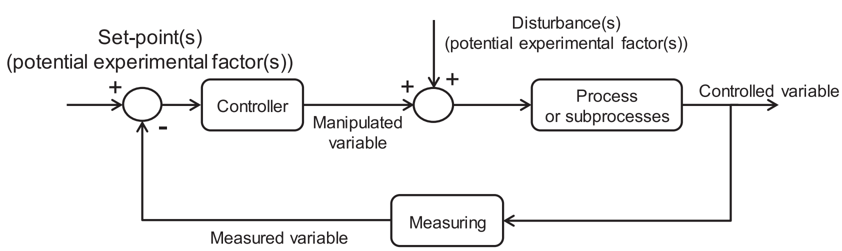

Figure 2. Schematic overview of a process under closed-loop control.

properly. This constitutes the first message we would like to convey in this article. That is, if the control action is ignored, the experimental factor changes will likely not affect the response (the controlled variable) significantly. An erroneous conclusion from the lack of detectable reaction would then be, depending on the effectiveness of the control action, that the factor is unimportant. However, if the presence of the controller is suspected or known, controlled variables may be used as responses primarily to test the presence and the effectiveness of the controllers. Manipulated variables may thus be considered as responses to study the impact of the experimental factors on the system and its dynamics due to the displacement of the potential effects from controlled to manipulated variables.

We classify experimental factors for processes operating under closed-loop control as (i) either a set of system inputs not involved in any control loop (should be viewed as disturbances in Figure 2) or (ii) the actual set-point values in the control loops. In the former scenario, both the manipulated and controlled variables can be used as experimental responses, while in the latter case more natural responses may be overall process performance indicators such as cost and/or product quality.

3. The Tennessee Eastman process simulator

Downs and Vogel[7] introduced the TE chemical process simulator for studying and developing engineering control design. The process is open loop unstable meaning that it will deviate and stop after a certain time period without any active control. With an appropriate control strategy, however, the process will remain stable. Several different control strategies for the TE process have been proposed; see for example McAvoy,[10] Lyman and Georgakis,[11] and Ricker.[12] The TE process has also been used as a test-bed for methodological development of multivariate statistical process monitoring.[8,13–16]

In the remainder of this section, we will describe some of the details of the TE process to facilitate the understanding of the experimental scenarios we use.

3.1. Process description

The TE process is a chemical process for which the components, kinetics, processing and operating conditions have been modified for proprietary reasons, see Downs and Vogel.[7] Following four irreversible and exothermic reactions, the process produces two liquid products from four gaseous reactants. With an additional byproduct and an inert product, eight components are present in the process. The process has five major unit operations: a *reactor*, a *product condenser*, a *vapor–liquid separator*, a *recycle compressor* and a *product stripper* as shown in a simplified process overview in Figure 3. A more detailed process map is given in the original reference.[7]

The physical inputs to the process consist of four gaseous streams, out of which three are fed to a reactor. After the reaction, the product mixture flows into a condenser, in which most of the gas is condensed. Some non-condensable components remain as vapors and the two phases are separated in the vapor–liquid separator. Vapor is partially recycled and purged together with the inert product and the byproduct. The product stripper separates remaining reactants from the products. The reactants are recycled, and the products exit the process from the stripper.

The TE process simulator has 12 manipulated variables (XMVs) and 41 measured variables (XMEASs). Out of 41 measured variables, 22 are measured directly while the remaining 19 variables can be calculated by the composition of the directly measured streams. In addition to XMVs and XMEASs, operating costs, production and product quality data are also recorded.

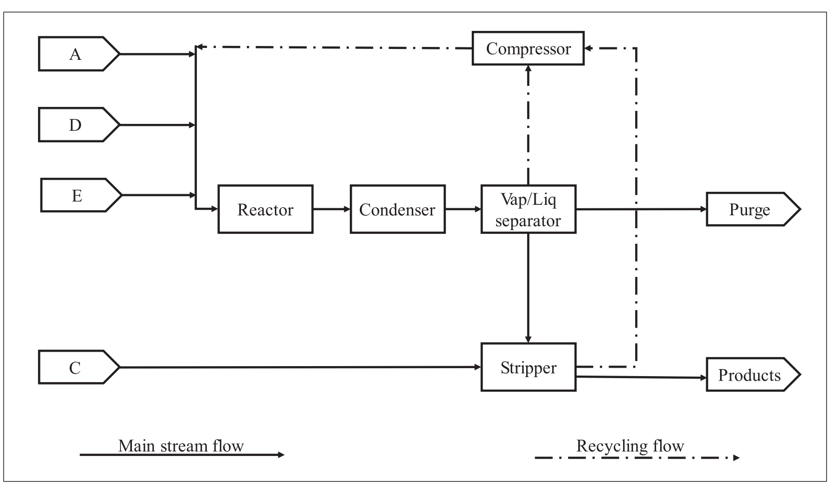

Figure 3. A schematic overview of the TE process.

The TE process has six different operating modes based on the production ratio of the two products and the production rate. Mode 1 is the most commonly used base case in research articles, which we also employ in this article. Five operating constraints need to be fulfilled to avoid process shutdown. There is also a possibility to activate 20 pre-set process disturbances (IDVs) during process operation. Downs and Vogel[7] provide more information on manipulated and measured variables, operating constraints, disturbances and the different operating modes.

3.2. *Implemented process control strategy*

A control strategy is a prerequisite for experimentation in the TE process because it is open loop unstable. Ricker[12] developed a decentralized control strategy for the TE process for improved performance, especially for the maximization of the production rate. The decentralized approach partitions the TE plant into 19 sub-units to each of which a controller is added. Tables I and II list the control loops, controlled variables, their set-points and manipulated variables. Note that we also provide XMV(i) and XMEAS(j); the i^{th} manipulated variable and the j^{th} measured variable given in Tables III and IV of the original article by Downs and Vogel[7] for ease of comparison. The manipulated variables listed with different codes, such as F_p, r_7 etc. come from the decentralized control strategy settings given in Ricker.[12]

We use a Matlab/Simulink decentralized control simulator (available at: http://depts.washington.edu/control/LARRY/TE/download.html#MATLAB_5x). In this configuration, all constraints are satisfied and the process can operate without undesired shutdowns. Moreover, the set-point values for some controlled variables and the values of inputs (XMVs) not involved in control loops may be varied, thereby allowing for experimentation.

The override loops 18 and 19 are exceptions to the control procedure described in Section 2. These control loops are only active when abnormal conditions occur that require an operating strategy change. Severe disturbances such as an introduction of the feed loss of A (IDV 6) activate the override loops. The production index F_p and the compressor recycle valve XMV(5) are not manipulated when the process operates without disturbances. All variables that can be manipulated except for the stripper steam valve XMV(9) and the agitator speed XMV(12) are involved in control loops in the decentralized control strategy. Consequently, XMV(9) and XMV(12) may be varied during experimentation and should then be viewed as disturbances in Figure 2.

3.3. *Chosen experimental scenarios in the TE process*

Two experimental scenarios in the TE process will illustrate experimentation in a process under closed-loop control. The first scenario will demonstrate an experiment when the system is disturbed by experimental factors. Input variables not involved in control loops can act as such disturbances and therefore be defined as experimental factors. The second scenario will demonstrate the use of the set-points of the control loops as experimental factors.

3.3.1. Scenario 1. The aim of this scenario is to demonstrate and visualize how experimental factor variation effects are distributed among the controlled and manipulated variables and how these effects can be analyzed. Recall that the stripper steam valve XMV(9)

Table I. Control loops for the decentralized control strategy (Ricker[12])

	Controlled variable		Manipulated variable	
Loop	Name	Code	Name	Code
1	A feed rate (stream 1)	XMEAS(1)	A feed flow	XMV(3)
2	D feed rate (stream 2)	XMEAS(2)	D feed flow	XMV(1)
3	E feed rate (stream 3)	XMEAS(3)	E feed flow	XMV(2)
4	C feed rate (stream 4)	XMEAS(4)	A and C feed flow	XMV(4)
5	Purge rate (stream 9)	XMEAS(10)	Purge valve	XMV(6)
6	Separator liquid rate (stream 10)	XMEAS(14)	Separator pot liquid flow	XMV(7)
7	Stripper liquid rate (stream 11)	XMEAS(17)	Stripper liquid product flow	XMV(8)
8	Production rate (stream 11)	XMEAS(41)	Production index	F_p
9	Stripper liquid level	XMEAS(15)	Ratio in loop 7	r_7
10	Separator liquid level	XMEAS(12)	Ratio in loop 6	r_6
11	Reactor liquid level	XMEAS(8)	Set-point of loop 17	s.p. 17
12	Reactor pressure	XMEAS(7)	Ratio in loop 5	r_5
13	Mol % G (stream 11)	XMEAS(40)	Adjustment to the molar feed rate of E	E_{adj}
14	Amount of A in reactor feed, y_A(stream 6)	XMEAS(6)	Ratio in loop 1	r_1
15	Amount of A + C in reactor feed, y_{AC} (stream 6)	XMEAS(6)	Sum of ratio in loop 1 and 4	$r_1 + r_4$
16	Reactor temperature	XMEAS(9)	Reactor cooling water flow	XMV(10)
17	Separator temperature	XMEAS(11)	Condenser cooling water flow	XMV(11)
18	Maximum reactor pressure	XMEAS(7)	Production index	F_p
19	Reactor level override	XMEAS(8)	Compressor recycle valve	XMV(5)

Table II. Set-point values in the control loops for the decentralized control strategy (Ricker[12])

Loop	Controlled variable	Set-point	
		Base case values	Units
1	A feed rate (stream 1)	0.2505	kscmh
2	D feed rate (stream 2)	3664.0	kg h^{-1}
3	E feed rate (stream 3)	4509.3	kg h^{-1}
4	C feed rate (stream 4)	9.3477	kscmh
5	Purge rate (stream 9)	0.3371	kscmh
6	Separator liquid rate (stream 10)	25.160	m^3 h^{-1}
7	Stripper liquid rate (stream 11)	22.949	m^3 h^{-1}
8	Production rate (stream 11)	100	%
9	Stripper liquid level	50	%
10	Separator liquid level	50	%
11	Reactor liquid level	75	%
12	Reactor pressure	2705	kPa
13	Mol % G (stream 11)	53.724	mol%
14	Amount of A in reactor feed, y_A (stream 6)	54.95	%
15	Amount of A + C in reactor feed, y_{AC} (stream 6)	58.57	%
16	Reactor temperature	120.40	°C
17	Separator temperature	80.109	°C
18	Maximum reactor pressure	2950	kPa
19	Reactor level override	95	%

Table III. Potential experimental factors in scenario 1. Input variables not involved in control loops

Variable name	Code	Base case value (%)	Low limit (%)	High limit (%)
Compressor recycle valve	XMV(5)	22.210	0	100
Stripper steam valve	XMV(9)	47.446	0	100
Agitator speed	XMV(12)	50.000	0	100

Table IV. Potential experimental factors of the TE process: set-point values of the control loops

Loop	Controlled variable	Base set-point
7	Stripper liquid rate (production)	22.949 m^3 h^{-1}
9	Stripper liquid level	50%
10	Separator liquid level	50%
11	Reactor liquid level	75%
12	Reactor pressure	2705 kPa
13	Mole % G	53.724 mol%
14	Amount of A in reactor feed (y_A)	54.95%
15	Amount of A + C in reactor feed (y_{AC})	58.57%
16	Reactor temperature	120.40 °C

and the agitator speed XMV(12) are the only two manipulated variables not involved in control loops. Moreover, if the process is run without introducing any of the pre-set disturbances (IDVs), the compressor recycle valve XMV(5) is not manipulated and can be considered as another possible experimental factor. Because the TE simulator is designed the way it is, these factors not involved in control loops can be seen as potential experimental factors (disturbances), and an experiment can be designed to evaluate their impact on the system. We would like to note that in a real process the experimental factors need not only come from a list of numeric input variables not involved in control loops but can rather be drawn from a variety of potential disturbances to the system, such as different raw materials, methods of operation etc. Our choice here is convenient because XMV(5, 9, and 12) can be changed rather easily in the simulation model.

Three experimental factors are thus available in this scenario. Response variables will be the controlled variables as well as the manipulated variables in the control loops (see Section 2). Table III presents base case values of XMV(5, 9 and 12) and their allowed ranges in operating Mode 1 of the TE process.

3.3.2. Scenario 2. The aim of this scenario in the TE process is to explore the set-points of the controllers to reveal their potential impact on the process operating cost. That is, to see causal relationships between the process' operating conditions and an important process performance indicator. By changing the set-points, the second experimental scenario indirectly uses the levels of the controlled variables as experimental factors. However, some of the set-points are actually controlled in a cascaded procedure based on directives generated by other controllers. Thus, only a subset of the controlled variables may be considered potential experimental factors. Table IV lists the controlled variables that may be used as potential experimental factors and their set-point values for operating Mode 1.

4. Scenario 1: design and analysis

This section and Section 5 through examples illustrate the two experimental scenarios explained above. We would like to clarify that the aim of these examples is not to show the 'best' experimental designs or analysis procedures but rather to illustrate issues related to experimentation in closed-loop operation.

4.1. A two-level factorial design

Scenario 1 involves a 2^2 randomized factorial design with three replicates with the aim of estimating location effects (main effects and interaction) of the stripper steam valve XMV(9) and of the agitator speed XMV(12) on controlled variables and associated manipulated variables. Control loops 9, 10, 11, 12 and 16 (see Table I) include constraints implemented for securing plant safety and adequate control actions to avoid shutdown.

The run-order of the experiments and the averages of the controlled and manipulated variables are given in Table V. The TE process was run for 36 h under normal operating conditions, i.e., the base case values for operating Mode 1, before starting the first experimental run. This 'warm-up phase' allows for the process to reach a steady-state condition before the manipulated variables are changed. Thereafter, every run lasted 50 h, and all 12 runs were run in sequence during continuous operation of the process. We did not apply any of the possible pre-set disturbances (IDVs) during experimentation. Including the warm-up phase, the entire experiment contained 636 h of simulated operation (real simulation time is only 115 s on a computer using an Intel® Core™ i5-4310 U processor running at 2.0 GHz with 16 GB of RAM.) The controlled and manipulated variables were sampled every 12 min.

Due to the process' continuous nature, the experimental factors and responses need to be viewed as time series. For example, Figure 4 illustrates the impact of the experimental factors on the controlled and manipulated variables in Loop 16 which controls the reactor temperature, XMEAS(9), by adjusting the reactor cooling water flow, XMV(10).

As seen in Figure 4, the experiment has a substantial impact on the manipulated variable – reactor cooling water flow, XMV(10). However, even though the levels of the experimental factors are changing, the controlled reactor temperature XMEAS(9) exhibits a random variation around its set-point value, indicating that the impact on this controlled variable is small or non-existent. A similar behavior has been observed also for loops 9, 10 and 11.

4.2. Statistical analysis

In the first scenario, the manipulated variables of loops 9, 10, 11, 12 and 16 are considered as the main response variables. A simple but reasonable way to analyze the experiments with time series responses is to ignore the time series aspect of the responses and to calculate the average value for each run in Table V, see Vanhatalo *et al.*[17]. Vanhatalo *et al.*[18] recommend removing apparent dynamic behavior at the beginning of each run. However, the initial observations are here included to investigate if the control loops are effective because the control action may not succeed to remove the impact on the controlled variable instantly. The run averages can be used to perform analysis of variance (ANOVA). Table VI presents a summary of the ANOVA based on the averages in Table V. The analysis was performed using the software Design-Expert® version 9.

Based on the high *p*-values for the controlled variables in Loops 11, 12 and 16, the results do not indicate that the experimental factors affect their related controlled variables. However, as revealed by the low *p*-values for the manipulated variables in Loops 12 and 16 in Table VI, the experimental factors affect process phenomena controlled by these loops. Furthermore, Loops 9 and 10 fail to remove the full impact of the experimental factor variation on the controlled variables as indicated by the low *p*-values on the controlled variables. There is no evidence that the experimental factors are affecting process phenomena controlled by Loop 12. Furthermore, the low *p*-value of the main effect of the stripper steam valve XMV(9) on the stripper liquid level in Loop 9, XMEAS(15), is explained by the inclusion of the transition time. The run averages are affected because the control action of Loop 9 is delayed.

4.3. Concluding remarks for scenario 1

When experimenting in a closed-loop system, the analyst should expect that the impact of the experimental factors could be partly or completely displaced from the controlled variables to manipulated variables. This is true despite using inputs not involved in control loops as experimental factors, if the experimental factors affect the phenomena controlled in the loops. However, as illustrated, the analysis may reveal potential ineffectiveness of the controllers to completely or instantly remove disturbances acting on controlled variables. We therefore recommend viewing the responses as two important and closely related groups: [1] controlled variables and [2] manipulated variables when analyzing an experiment in a closed-loop system as illustrated above.

Table V. Run order and averages of the observations for the manipulated and controlled variables in the 2^2 factorial experiment

Run	Experimental factors		Manipulated variables (XMVs)					Controlled variables (XMEASs)				
	XMV(9) %	XMV(12) %	Loop 9 r_7	Loop 10 r_6	Loop 11 s.p. 17	Loop 12 r_5	Loop 16 XMV(10)	Loop 9 $j = 15$	Loop 10 $j = 12$	Loop 11 $j = 8$	Loop 12 $j = 7$	Loop 16 $j = 9$
1	40	40	0.2273	0.2489	80.3247	0.0034	42.5970	49.97	49.97	74.98	2705.03	120.40
2	60	60	0.2281	0.2500	80.2536	0.0034	39.7900	50.34	50.09	74.94	2705.04	120.40
3	60	40	0.2281	0.2500	80.2754	0.0034	42.4825	50.06	49.97	75.01	2705.03	120.40
4	40	60	0.2272	0.2489	80.3755	0.0034	39.8809	49.62	50.00	75.02	2704.99	120.40
5	60	60	0.2281	0.2501	80.2909	0.0033	39.7862	50.37	50.19	74.99	2704.95	120.40
6	40	40	0.2272	0.2488	80.3795	0.0033	42.6090	49.62	50.02	74.98	2705.02	120.40
7	40	40	0.2272	0.2489	80.3306	0.0033	42.5969	49.92	50.04	74.99	2705.00	120.40
8	60	40	0.2281	0.2501	80.2541	0.0033	42.4961	50.16	50.04	74.98	2704.99	120.40
9	60	40	0.2281	0.2501	80.2469	0.0033	42.5057	49.98	49.97	75.02	2705.01	120.40
10	40	60	0.2272	0.2487	80.3858	0.0033	39.8842	49.81	49.93	74.99	2705.05	120.40
11	40	60	0.2272	0.2489	80.3315	0.0033	39.8970	49.97	49.83	74.96	2704.95	120.40
12	60	60	0.2281	0.2500	80.2335	0.0033	39.8065	50.36	50.00	75.06	2705.01	120.40

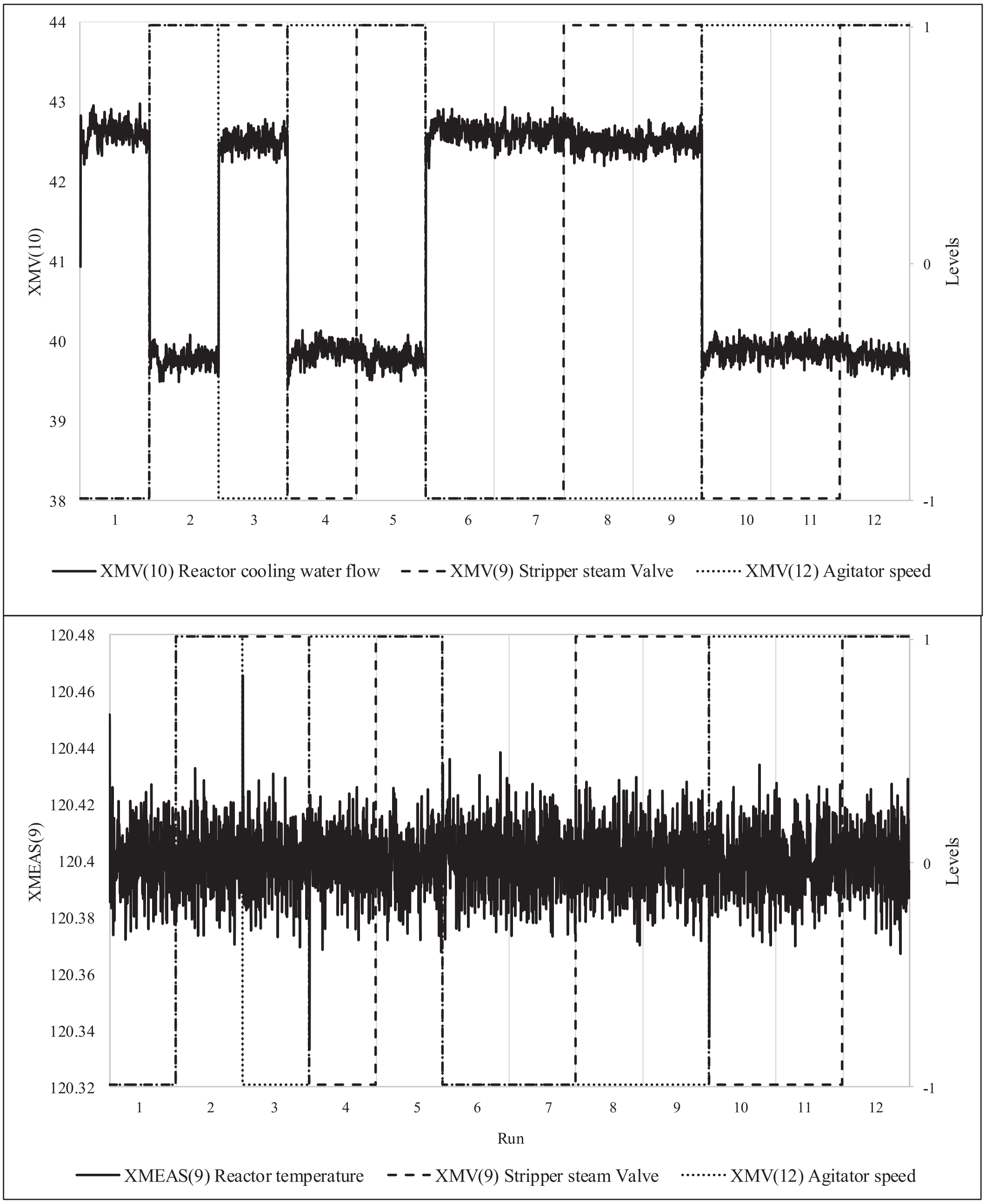

Figure 4. Overview of experimental factors' impact on variables related to control loop 16. The manipulated variable, XMV(10), is given in the top chart and controlled variable, XMEAS(9), in the bottom chart. The levels, in coded units, of the experimental factors XMV(9) and XMV(12) are superimposed on the plots. The duration of each experiment is 50 h.

5. Scenario 2: design and analysis

The second scenario illustrates a different way of running experiments in closed-loop controlled processes. Now, we consider the set-points of the control loops as experimental factors. Our major concern is no longer to reveal cause and effect relationships between inputs and important measured variables in the process. These should have been identified already in the engineering control design phase. Instead, we are exploring the set-points of the controllers to see causal relationships between the process operating conditions and process performance indicators with the aim of optimizing the process.

5.1. A screening experiment

In this case, we focus on the process operating cost as an important response. We have nine possible set-points to change (see Table IV), and we wish to test their impact on the process operating cost using a two-step sequential experiment. The starting point

Table VI. The *p*-values of the estimated effects based on ANOVA. Cells with bold text indicate the significant effects based on a significance level of 5%

Main effects and interaction	Manipulated variables					Controlled variables XMEAS(*j*)				
	Loop 9	Loop 10	Loop 11	Loop 12	Loop 16	Loop 9	Loop 10	Loop 11	Loop 12	Loop 16
	r_7	r_6	sp17	r_5	XMV(10)	$j = 15$	$j = 12$	$j = 8$	$j = 7$	$j = 9$
XMV(9)	**<0.0001**	**<0.0001**	**<0.0002**	0.7011	**<0.0001**	**0.001**	0.0937	0.4997	0.9031	0.9588
XMV(12)	0.9958	0.3744	0.5351	0.4029	**<0.0001**	0.1414	0.8736	0.8567	0.4668	0.1363
*XMV(9)*XMV(12)*	0.2754	0.9865	0.5564	0.9997	0.2759	0.0692	**0.0412**	0.7048	0.8390	0.6849

is a 2_{III}^{9-5} fully randomized fractional factorial design with four additional center points. This resolution III design is then followed by a full fold-over in a new block to entangle some aliased effects. The final design, i.e., the original plus the fold-over, is of resolution IV.

Some factor setting combinations will invoke a process shutdown and some shutdown limits are also given in the Downs and Vogel[7] paper. The base case value of each factor (rounded to the nearest integer) was chosen as either the low or high factor level in the design. The other level of each variable was defined by trial and error by either adding to or subtracting from the base case value while trying to keep the process from shutting down. Table VII provides the low and high levels of each experimental factor (set-point) used in the experiment.

Furthermore, we chose to keep XMVs (5, 9 and 12) fixed at their base case values given in Table III during the experiment because they are not involved in the loops but do affect the process behavior.

A 'warm-up phase' of 36 h was once again used before the start of the first run of the experiment. During this phase, the experimental factors (set-points) were fixed to their base case values for operating Mode 1. The 40 runs of the experiment are given in Table VIII. Each experimental run lasted 50 h. Including the warm-up phase, the entire experiment contained 2036 h of operation (simulation time was 147 s for all runs). From the TE simulator, the process operating cost ($/h) can be extracted, and we have the operating cost for every 12 min. Figure 5 illustrates the impact of the experimental factors on the process operating cost during the first three experiments in run order.

Table VII. Low and high level of the set-points used as experimental factors

Loop	Controlled variable	Base set-point	Low level	High level
7	Stripper liquid rate (production)	22.949 $m^3 h^{-1}$	21 $m^3 h^{-1}$	23 $m^3 h^{-1}$
9	Stripper liquid level	50%	50%	60%
10	Separator liquid level	50%	35%	50%
11	Reactor liquid level	75%	70%	75%
12	Reactor pressure	2705 kPa	2600 kPa	2705 kPa
13	Mole % G	53.724 mol%	54 mol%	62 mol%
14	Amount of A in reactor feed (y_A)	54.95%	55%	65%
15	Amount of A + C in reactor feed (y_{AC})	58.57%	50%	59%
16	Reactor temperature	120.40 °C	120 °C	127 °C

Table VIII. Run order, standard order of the runs and average operating cost both before and after removal of transition time at the beginning of each run

Block 1: 2_{III}^{9-5} experimental design				Block 2: Full fold-over			
Run order	*Standard order*	*Operating cost ($/h)*	*Operating cost ($/h) (after removing transition time)*	*Run order*	*Standard order*	*Operating cost ($/h)*	*Operating cost ($/h) (after removing transition time)*
1	14	201.11	201.68	21	38	139.46	130.84
2	2	156.51	154.51	22	26	130.55	131.72
3	9	148.60	143.56	23	34	152.75	146.08
4	4	127.37	140.00	24	27	156.25	157.61
5	6	185.37	172.01	25	35	182.89	170.58
6	20	124.19	129.89	26	22	125.28	126.76
7	1	139.87	141.24	27	30	175.37	157.19
8	17	133.27	131.09	28	39	120.70	131.02
9	11	123.56	129.74	29	33	151.78	150.66
10	12	255.76	215.15	30	24	166.46	155.20
11	8	175.52	187.61	31	29	129.43	142.91
12	16	164.44	160.05	32	31	186.93	167.84
13	18	127.84	130.15	33	28	166.30	167.94
14	15	147.23	142.59	34	36	164.98	165.41
15	19	130.64	132.81	35	37	128.14	132.72
16	5	104.70	109.27	36	21	145.67	140.70
17	13	181.27	161.61	37	32	104.34	115.46
18	3	128.85	127.87	38	23	174.01	166.69
19	7	182.26	177.45	39	25	213.02	198.88
20	10	117.49	127.62	40	40	127.23	135.06

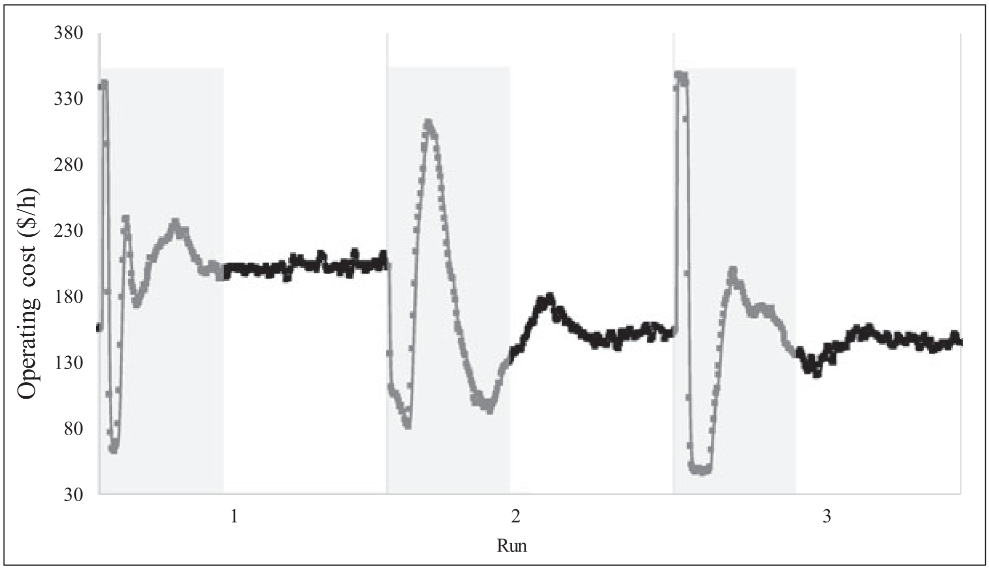

Figure 5. The operating cost during the first three runs of the experiment. Note the dynamic behavior of the response during the first part of each run. The shaded areas highlight the removed observations before calculating the run averages. The duration of each experiment is 50 h.

5.2. Statistical analysis

The aim of the experiment is to find set-points which reduce the long-term operating cost. In contrast to scenario 1, it makes sense to remove transition time from the runs and then use the remaining observations to calculate run averages. To keep the observations during the transition time in the calculation of run averages will lead to an underestimation of the location effect of the factors and interactions, see Vanhatalo et al.[17] The process operating cost exhibits some transition time before reaching the steady state as illustrated in Figure 5. A visual inspection of the operating cost reveals that 24 h can be considered as a reasonable transition time (grey shaded area in Figure 5), and thus the observations obtained during the first 24 h of all runs were removed before calculating the run averages, see Table VIII.

Table IX presents an ANOVA table of the 40-run experimental design in Table VIII based on a significance level of 5%. We have also repeated the analysis including the transition time. The results of that analysis are not reported in this article, but with the transition time included, the same main effects turn out to be active, but the significant interaction effects differ. As seen in Table IX, seven main effects and eight two-factor interaction alias strings are active (interactions of order three or higher are ignored). It is perhaps not surprising that most factors affect the operating cost because control loops aim to control important process phenomena which tend to affect the production cost. Moreover, the interconnectedness of the different control loops is demonstrated by the many significant interactions.

Note that the curvature test is significant and that the model exhibits significant lack of fit, suggesting that a higher order model is appropriate. The fitted model in Table IX is thus ill-suited for optimization and prediction but provides a starting point for future response surface experimentation. The many significant two-factor interaction alias strings would need further investigation to decide which among the aliases are actually active. However, as we mentioned earlier, the main purpose of this article is not necessarily to provide an optimization procedure on a simulated process but rather to draw attention to possibilities and pitfalls in experimentation under closed-loop operation. Hence, for demonstration purposes, we simply assume that the first interactions of the interaction strings in Table IX are the important ones, ignoring the interactions in brackets. We proceed to use the estimated model to provide suggested factor settings for the lowest operating cost within the experimental region. In this case, the lowest cost will be at a corner point on the multidimensional hyperplane. The settings of the factors and the predicted operating cost at this point (104.5 $/h) are provided in Table X. The significant curvature, the lack of fit tests and the R^2 for prediction indicate that the predictive ability of the model is poor. A confirmation run in the TE process simulator using the suggested factors settings gives the long-term average operating cost 109.1 $/h. The 4.6 $/h discrepancy between the predicted cost and the confirmation run is likely due to the models' poor predictive ability. Nevertheless, this rough analysis provides a significant improvement of the process operating cost. A simulation of the process keeping the factors settings at the base set-points values given in Table VII gives a long-term average operating cost of 170.2 $/h. Hence, running the process at the suggested factors settings leads to a substantial cost reduction of 61.1 $/h. Further reduction of the operating cost is likely possible but outside the scope of this article.

5.3. Concluding remarks for scenario 2

The second scenario illustrates how designed experiments can be used to improve process performance indicators using the set-points of variables controlled in closed-loop. This scenario also exemplifies the importance of considering, and here removing, the transition time during analysis. We want to point out that the set-points of the controllers in this example and in real life in general affect important process operating conditions. The experimenter should therefore expect that improper choices of factor levels of the

Table IX. ANOVA and estimated effects based on the averages of the response after removing the transition time. The model includes only terms significant at 5% level. Aliased two-factor interaction aliases that based on the heredity principle are less likely given in italic text within brackets. The control loop numbers are indicated by (#) in the factor names

Source	Sum of squares	*df*	Mean square	*F* value	Prob > *F*	Estimated standardized effects
Block	15.11	1	15.11			
Model	18 431.18	15	1228.75	83.97	<0.0001	
A: #7—Production	3946.30	1	3946.30	269.68	<0.0001	11.11
D: #11—Reactor level	321.10	1	321.10	21.94	0.0001	3.17
E: #12—Reactor pressure	3131.68	1	3131.68	214.01	<0.0001	−9.89
F: #13—Mole %G	4085.75	1	4085.75	279.21	<0.0001	−11.30
G: #14—y_A	443.48	1	443.48	30.31	<0.0001	−3.72
H: #15—y_{AC}	2444.72	1	2444.72	167.07	<0.0001	8.74
J: #16—Reactor temp	126.25	1	126.25	8.63	0.0076	−1.99
AD (*BH CG* FG)	124.36	1	124.36	8.50	0.0080	1.97
AF (*BG CH* DE)	207.73	1	207.73	14.20	0.0011	−2.55
AG (*BF CD* EH)	78.98	1	78.98	5.40	0.0298	−1.57
AH (*BD CF* EG)	151.98	1	151.98	10.39	0.0039	−2.18
AJ	532.42	1	532.42	36.38	<0.0001	−4.08
FJ	282.93	1	282.93	19.34	0.0002	2.97
GJ	619.92	1	619.92	42.36	<0.0001	4.40
HJ	1933.59	1	1933.59	132.14	<0.0001	7.77
Curvature	3415.43	1	3415.43	233.40	<0.0001	
Residual	321.93	22	14.63			
Lack of Fit	305.16	16	19.07	6.82	0.0129	
Pure Error	16.77	6	2.80			
Cor Total	22 183.66	39				
				R^2		83.1%
				Adjusted R^2		72.1%
				R^2 prediction		67.2%

Table X. Suggested settings of the set-points of the control loops to provide the lowest operating cost of the estimated model within the experimental region

Loop	Set-point	Suggested setting
7	Stripper liquid rate (production)	21 $m^3 h^{-1}$
9	Stripper liquid level	Not in model, use base case
10	Separator liquid level	Not in model, use base case
11	Reactor liquid level	70%
12	Reactor pressure	2705 kPa
13	Mole % G	62 mol%
14	Amount of A in reactor feed (y_A)	65%
15	Amount of A + C in reactor feed (y_{AC})	50%
16	Reactor temperature	120 °C
Resulting predicted process operating cost:		104.5 $/h

set-points may lead to unexpected process behavior or even shutdown. Special care should be taken in choosing the levels because the window of operability may be irregular or unknown.

6. Conclusion and discussion

This article explores important issues in designing and analyzing experiments in the presence of engineering process control. The closed-loop operation increases process complexity and influences the strategy of experimentation. Two experimental scenarios based on the TE process simulator are used to answer the questions *why* and *how* to conduct and analyze experiments in closed-loop systems.

Even though we have prior experience with experiments lasting several weeks in continuous processes, the 2038 h of experimentation we use in our examples may admittedly be considered unrealistically long in practice. This is, however, beside the point because the examples we provide are for demonstration purposes, and we did not necessarily focus on shortening the duration of the experiments.

The first experimental scenario illustrates how the experimental factors not directly involved in control loops impact the closed-loop system and how the controllers affect the analysis. The controllers adjust manipulated variables to limit or eliminate the experimental factor effects on the controlled variable(s). We note that this will only occur if the experimental factors affect phenomena/variables governed by the closed-loop system. The effect on the controlled variables is partly or fully transferred to the manipulated variables depending on the effectiveness of the controllers. Hence, both the controlled and manipulated variables should be used as responses. Analyzing the effects of experimental factors on controlled variables may give important information about the effectiveness of the engineering process control. The effects on the manipulated variables instead reveal whether the experimental factors affect important process behavior.

In the second scenario, the experimental factors are the set-points of the controlled variables. The set-points are target values for the controlled variables and are typically closely tied to important process operating conditions. A level change of the set-points can therefore be considered equivalent to shifting the location of the process. Overall process performance indicators such as operating cost or product quality may then be suitable responses.

Using two scenarios we have illustrated that DoE can generate knowledge and aid process improvement in closed-loop systems. More specifically, DoE can be used to study:

- if the engineering process control is efficient and cost effective;
- if experimental factors affect important process phenomena; and
- how controlled variable set-points affect important process performance indicators.

We believe simulation software like the TE process offer great opportunities for methodology development in experimentation in closed-loop systems. In this article, we simply provide some basic ideas and approaches, but more research is needed for further development of experimentation and analysis methods for better process understanding and optimization in closed-loop systems.

Acknowledgments

The authors gratefully acknowledge the financial support from the Swedish Research Council under grant number 340-2013-5108.

References

1. Box GEP, MacGregor JF. The analysis of closed-loop dynamic-stochastic systems. *Technometrics* 1974; **16**(3):391–398.
2. Box GEP, MacGregor JF. Parameter estimation with closed-loop operating data. *Technometrics* 1976; **18**(4):371–380.
3. Ljung L. System Identification: Theory for the User (2nd. edn). Prentice-Hall: Upper Saddle River, NJ, 1999.
4. Jansson H. Experiment design with applications in identification for control. Department of Signals, Sensors, and Systems, Royal Institute of Technology (KTH), Stockholm, Sweden, 2004. Doctoral Thesis.
5. Vanhatalo E, Bergquist B. Special considerations when planning experiments in a continuous process. *Quality Engineering* 2007; **19**(3):155–169. doi:10.1080/08982110701474100.
6. Hild C, Sanders D, Cooper T. Six Sigma* on continuous processes: how and why it differs. *Quality Engineering* 2001; **13**(1):1–9. doi:10.1080/08982110108918618.
7. Downs JJ, Vogel EF. A plant wide industrial process control problem. *Computers & Chemical Engineering* 1993; **17**(3):245–255.
8. Kruger U, Zhou Y, Irwin GW. Improved principal component monitoring of large-scale processes. *Journal of Process Control* 2004; **14**(8):879–888. doi:10.1016/j.jprocont.2004.02.002.
9. Montgomery DC. Design and Analysis of Experiments (8th. edn). Wiley: New York, 2012.
10. McAvoy TJ, Ye N. Base control for the Tennessee Eastman problem. *Computers & Chemical Engineering* 1994; **18**(5):383–413. doi:10.1016/0098-1354(94)88019-0.
11. Lyman PR, Georgakis C. Plant-wide control of the Tennessee Eastman problem. *Computers & Chemical Engineering* 1995; **19**(3):321–331. doi:10.1016/0098-1354(94)00057-U.
12. Ricker LN. Decentralized control of the Tennessee Eastman challenge process. *Journal of Process Control* 1996; **6**(4):205–221. doi:10.1016/0959-1524(96)00031-5.
13. Ku W, Storer RH, Georgakis C. Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems* 1995; **30**(1):179–196. doi:10.1016/0169-7439(95)00076-3.
14. Lee G, Han C, Yoon ES. Multiple-fault diagnosis of the Tennessee Eastman process based on system decomposition and dynamic PLS. *Industrial & Engineering Chemistry Research* 2004; **43**(25):8037–8048. doi:10.1021/ie049624u.
15. Hsu C, Chen M, Chen L. A novel process monitoring approach with dynamic independent component analysis. *Control Engineering Practice* 2010; **18**:242–253. doi:10.1016/j.conengprac.2009.11.002.
16. Liu K, Fei Z, Yue B, Liang J, Lin H. Adaptive sparse principal component analysis for enhanced process monitoring and fault isolation. *Chemometrics and Intelligent Laboratory Systems* 2015; **146**:426–436. doi:10.1016/j.chemolab.2015.06.014.
17. Vanhatalo E, Bergquist B, Vännman K. Towards improved analysis methods for two-level factorial experiment with time series responses. *Quality and Reliability Engineering International* 2013; **29**(5):725–741. doi:10.1002/qre.1424.

18. Vanhatalo E, Kvarnström B, Bergquist B, Vännman K. A method to determine transition time for experiments in dynamic processes. *Quality Engineering* 2010; **23**(1):30–45. doi:10.1080/08982112.2010.495099.

Authors' biographies

Francesca Capaci is a doctoral candidate in Quality Technology at the Department of Business Administration and Social Sciences at Luleå University of Technology (LTU) in Sweden. She received her Master's degree in Industrial Engineering and Management at Universitá Degli Studi di Palermo (Italy) in 2014. Her research area is focused around experimental strategies and statistical process control for continuous processes. Since 2014, she has also been teaching Design of Experiment and Statistical Process Control at LTU. She is a member of the European Network for Business and Industrial Statistics (ENBIS).

Bjarne Bergquist is Professor and Chair of Quality Technology and Management at Luleå University of Technology, Sweden. His current research interests are focused on the use of statistically based methods and tools for improvements, particularly in the continuous process industries, as well as on statistically based railway maintenance. He has a background in Materials Science but has lately mainly published the field of operations management, quality technology and quality improvements.

Murat Kulahci is a professor in the Department of Business Administration, Technology and Social Sciences at Luleå University of Technology in Sweden and an associate professor in the Department of Applied Mathematics and Computer Science, at the Technical University of Denmark. His research focuses on design of physical and computer experiments, statistical process control, time series analysis and forecasting, and financial engineering. He has presented his work in international conferences, and published over 70 articles in archival journals. He is the co-author of two books on time series analysis and forecasting.

Erik Vanhatalo is an assistant professor and senior lecturer of quality technology and management at Luleå University of Technology (LTU). He holds an MSc degree in industrial and management engineering from LTU and a PhD in quality technology and management from LTU. His current research is focused on the use of statistical process control, experimental design, time series analysis and multivariate statistical methods, especially in process industry. He is a member of the European Network for Business and Industrial Statistics (ENBIS).